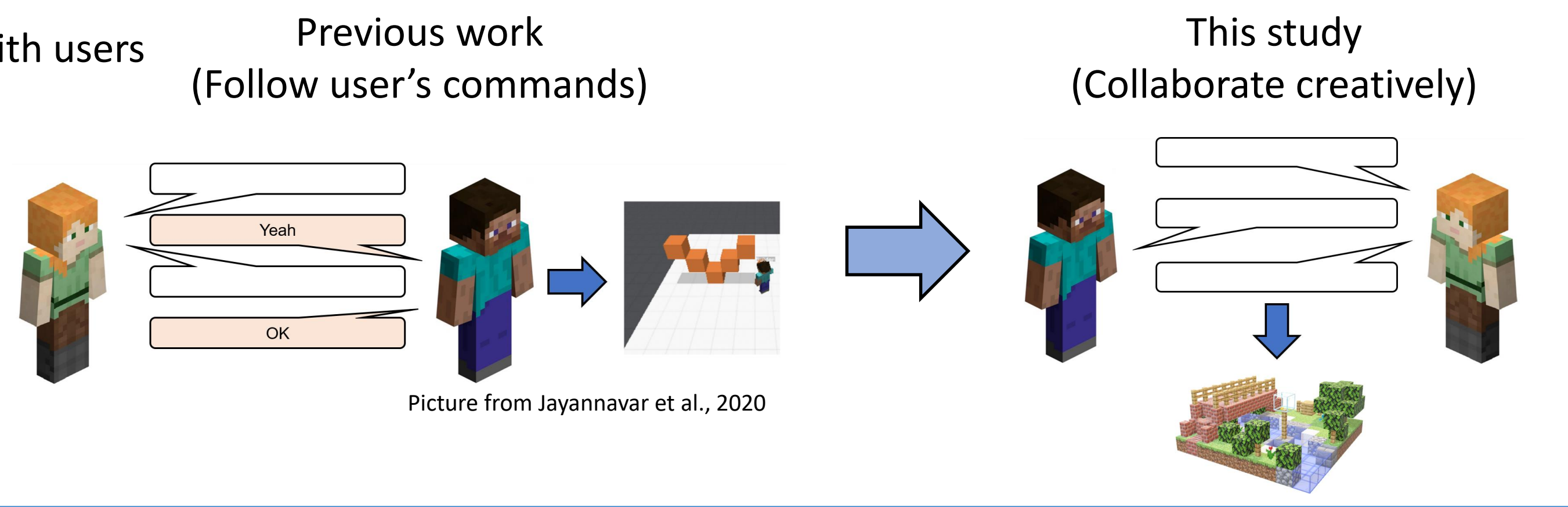


Overview

- Dialogue systems become part of our daily lives
- Previous systems have only followed user's commands, not been able to create a new value by collaborating with users
- It is essential that systems and users can work together on creative collaborative work
- Devised the Collaborative Garden Task in which two workers work together to create a garden in Minecraft
- Collected 500 human-human dialogues in Japanese
- Discovered two essential skills for dialogue in creative collaborative work
 - To be able to perform more processes to ask for and agree on suggestions between workers
 - To be able to agree on a particular image of the final product, then discussing changes and details



Collaborative Garden Task Corpus

- Collaborative Garden Task
 - Two workers **work together to create an original and beautiful garden**
 - Workers interacted via text chat (in Japanese) with their partner and manipulated blocks
 - Time limit for work was 20 min.
 - After work, workers answered a questionnaire about their satisfaction with the dialogue and the garden
 - Workers were paid approx. five dollars for one dialogue
- Dialogue Example



ID	S	Utterance (translated to English by authors)
1	A	Do you want to make something?
2	B	I want to make a roof like a wisteria trellis.
3	A	Sounds good!
4	A	Try making it in the middle!
5	B	I will make a path down the middle.
6	B	I also want to make something on both sides.
7	A	This side, is this okay?
8	B	It's good!
9	B	Can I try to put a chair at the two trees?
10	A	Go ahead! If it's in the way, please remove it!
11	B	Thank you.
12	B	Do you know how to make a nice chair and desk?
13	A	I cannot think of any... I wish I had a half block or something...
14	B	Yeah...
15	A	So big!
16	B	I wonder if it is too much to ask for a table with just one of these...
17	A	No... that's okay!
18	B	Great!



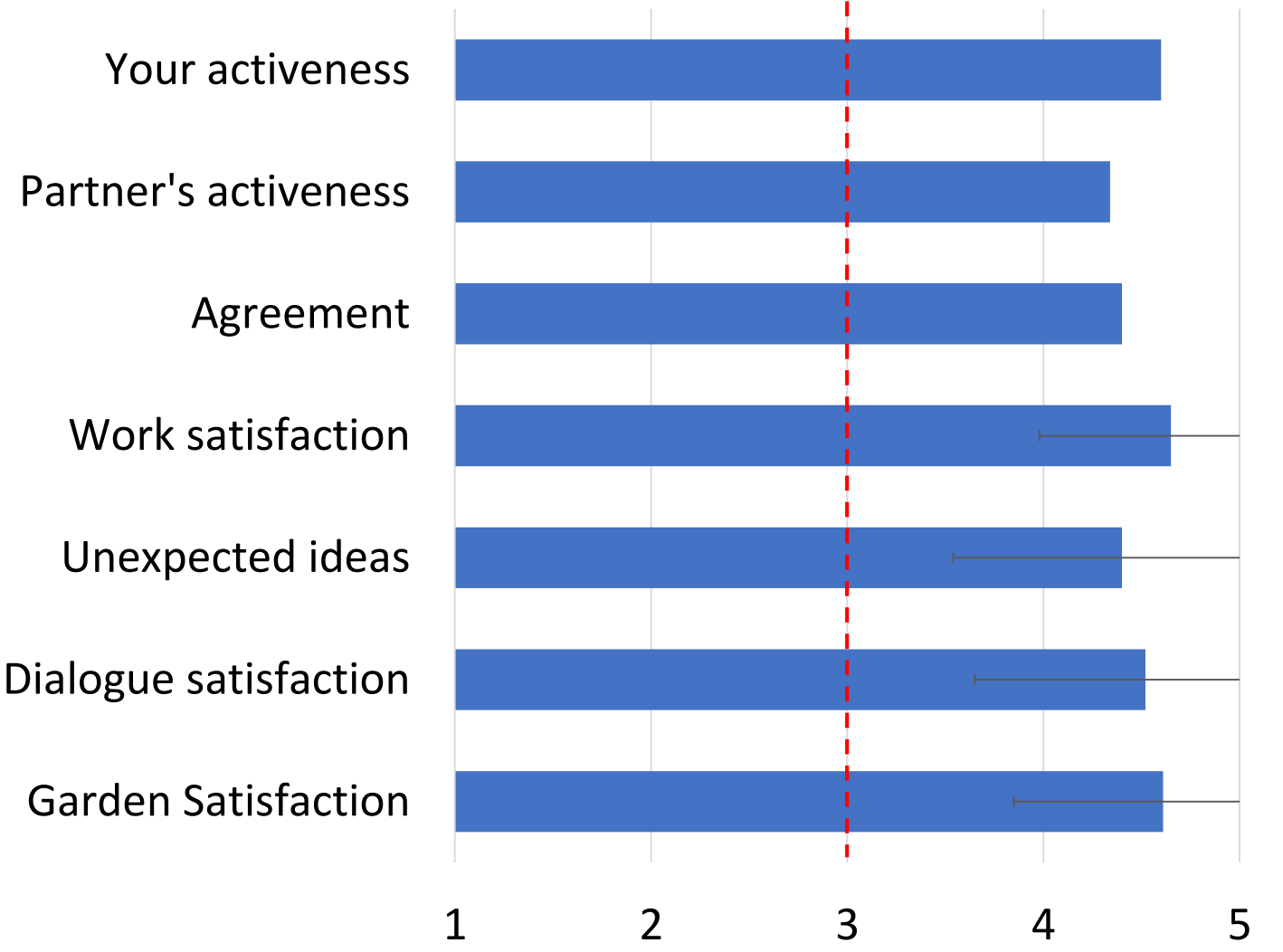
Corpus Statistics

	Total	Mean		Total	Mean
Number of workers	79		Number of placements	182,355	182.4 / worker
Number of pairs	500		Number of removals	131,552	131.6 / worker
Number of dialogues	500		Move distances		1,899.5 / worker
Number of utterances	16,221	32.4 / dialogue	Number of blocks		201.8 / garden
Number of characters	224,885	13.9 / utterance	Number of block types		11.6 / garden
Number of words	131,937	8.13 / utterance			
Number of different words		105.8 / dialogue			

- Enough utterances for statistical learning
- Workers repeatedly tried to place and remove blocks

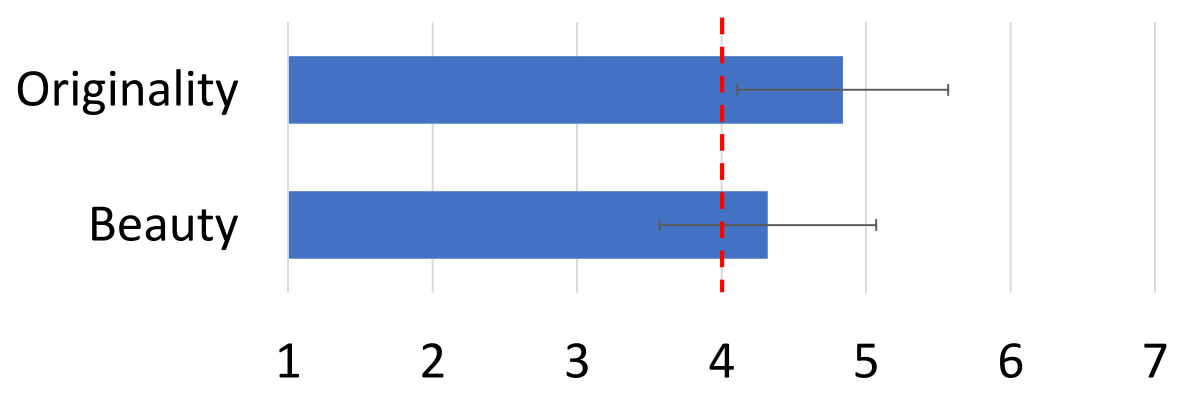
Subjective Evaluations

- Post-work questionnaire (five-point Likert scale)
- All items were highly rated, indicating **high-quality of the collected collaborative work**



Third-Person Evaluations

- Objectively assess the quality of the collected collaboration
- Crowd-workers evaluated the created gardens from the view point of originality and beauty (seven-point Likert scale)
- All items were relatively highly rated, indicating **good quality of the collected collaborative work**



Analysis of Dialogue in Collaboration

Mining Frequent Word 4-grams

- Methods
 - Classified 500 dialogues into the top 20% and the bottom 80% on evaluations scores
 - Focused on word 4-grams that contain at least one content word
 - Conducted Fisher's exact test on a 2 × 2 cross table to **identify characteristic 4-grams in high-quality dialogues**
 - One axis indicating whether a dialogue included a specific word 4-gram
 - Other indicating whether a dialogue was among the high-quality dialogues

Results

- Two main types of expressions were characteristics of high-quality dialogues
 - Asking for a partner's agreement
 - Showing positive appreciation toward the partner

Classification criteria	Expression (translated to English by authors)	Example utterance
Satisfaction (Total scores of the subjective evaluations)	would you like ...? how about ...?, let's try I think we could try	Bridge? Well, let's try to create a river underneath!
	that's right great!, good!	Yeah!
	thank you!, thank you it looks, it has a feel	It looks moderately calm , nice!
Originality	how about ...?	How about a private beach?
	it looks like	Now it looks like a shrine.
Beauty	would you like ...? let's try, I would like to	Would you like to place flowers to make it more colorful?
	looks nice	Looks nice!

It is essential to **perform more processes to ask for and agree on suggestions between workers**

Future Work

- Train an encoder-decoder model that can output user utterances on the basis of the garden's state and the dialogue history
- Work on developing a system that can creatively collaborate with users

Analysis of Patterns of Collaboration Process

Methods

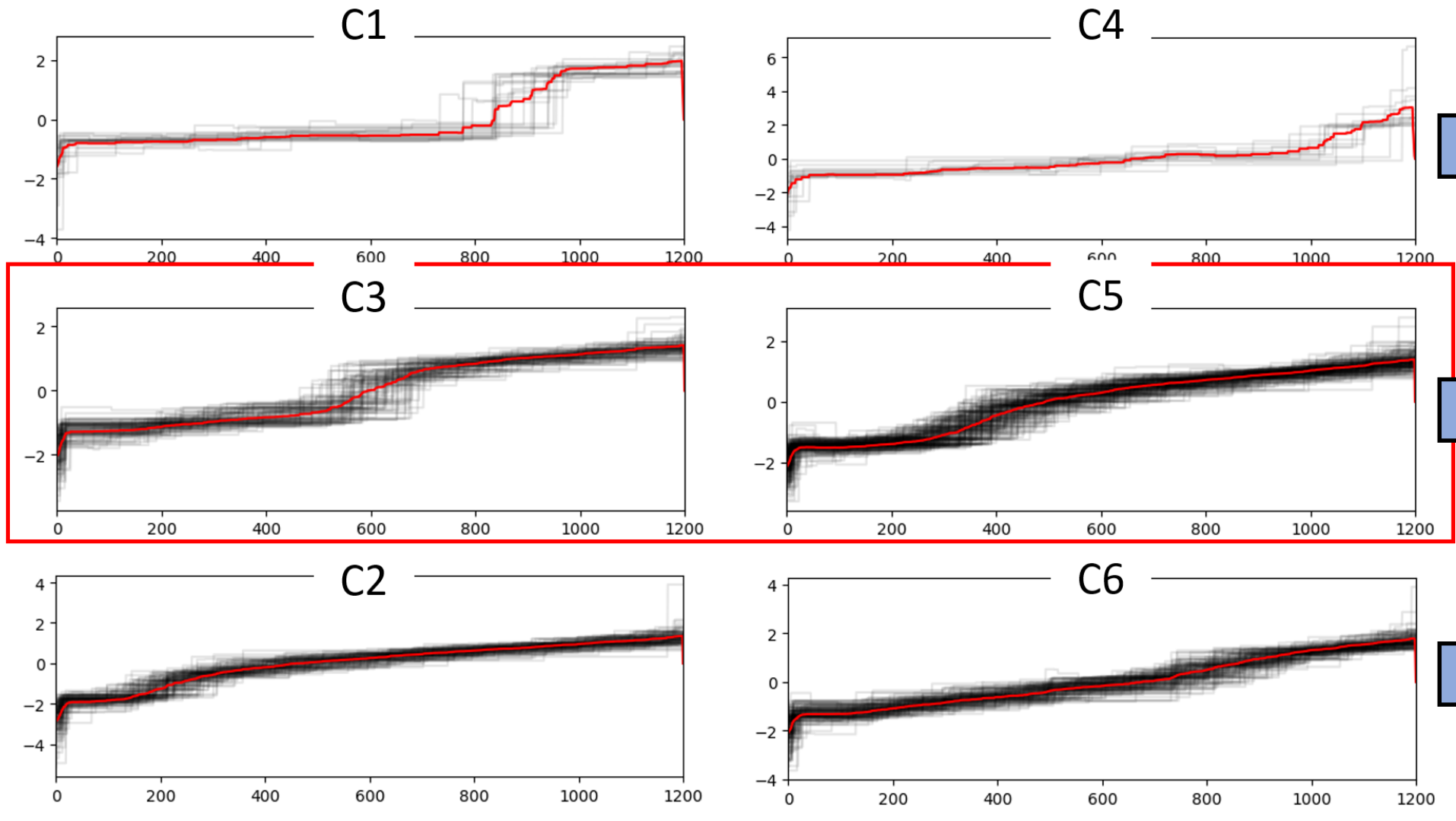
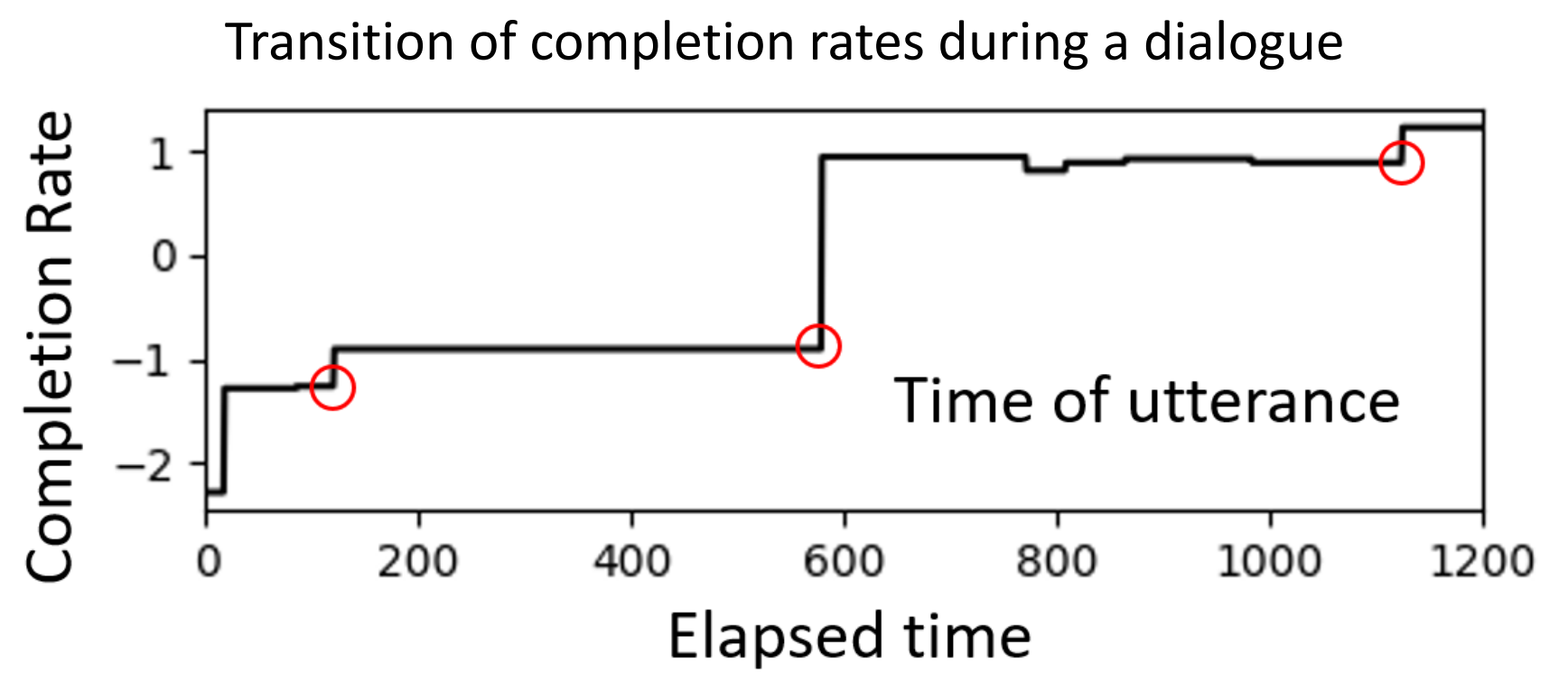
- Plotted the time-series data of the completion rate at the time of an utterance

$$\text{Completion Rate} = \frac{|Blocks_t \cap Blocks_{last}|}{|Blocks_{last}|}$$

- Conducted a cluster analysis using k-Shape

Results

- Obtained six clusters



- Workers worked in silence after deciding what to make in the early phase
- Workers created the garden to some degree and then discussed the details while looking at created objects
- Workers discussed and agreed on the garden step by step

Further Analysis

- Conducted Fisher's exact test on a 2 × 2 cross table to **identify which cluster was associated with high-quality dialogues**
 - One axis indicating whether a dialogue was in a specific cluster
 - Other indicating whether a dialogue was among the high-quality dialogues

Results

- High-quality dialogues appeared significantly ($p < 0.05$) more frequently in C3 and C5**

It is essential to **agree on a particular image of the final product in the early phase and then discussing the details to complete the garden**